AD-A048 887 CARNEGIE-MELLON UNIV PITTSBURGH PA MANAGEMENT SCIENC--ETC F/G 12/1
BENDERS'S METHOD REVISITED.(U)
MAY 77 E BALAS, C BERGTHALLER N00014-75-C-0621
UNCLASSIFIED MSRR-401 NL

1 OF 2
AD
A048887

~~END~~ CONT.
DATE
FILMED
2-78
DDC

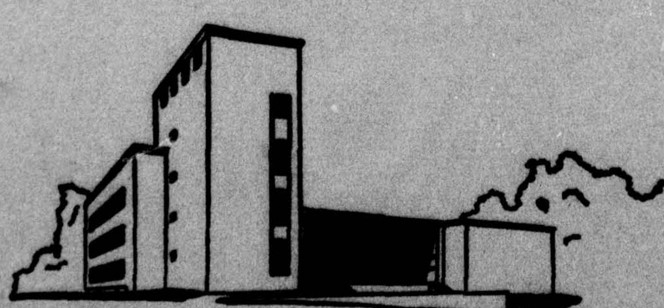
1

Carnegie-Mellon University

PITTSBURGH, PENNSYLVANIA 15213

GRADUATE SCHOOL OF INDUSTRIAL ADMINISTRATION

WILLIAM LARIMER MELLON, FOUNDER



W.P.#42-76-77 (1)

Management Sciences Research Report No. 401

BENDERS'S METHOD REVISITED

by

Egon Balas and Christian Bergthaller*

May 1977

Revised October 1977

This report was prepared as part of the activities of the Management Sciences Research Group, Carnegie-Mellon University, under Grant MPS73-08534 A02 of the National Science Foundation and Contract N00014-75-C-0621 NR 047-048 with the U.S. Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

Management Science Research Group
Graduate School of Industrial Administration
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

*Institute of Metrology, Bucharest, Romania (Visiting Exchange Scientist, Carnegie-Mellon University, September-December 1975)

(See 1473)

ABSTRACT

The master problem in Benders's partitioning method is an integer program with a very large number of constraints, each of which is usually generated by solving the integer program with the constraints generated earlier. Computational experience shows that the subset B of those constraints of the master problem that are satisfied with equality at the linear programming optimum often play a crucial role in determining the integer optimum, in the sense that only a few of the remaining inequalities are needed. We characterize this subset B of inequalities. Though the best upper bound (often attained) on the cardinality of B is 2^p, where p is the number of integer-constrained variables that are basic at the linear programming optimum, none of the inequalities in B is implied by the remaining inequalities of the master problem. We then give an efficient procedure for generating an appropriate subset of the inequalities in B, which leads to a considerably improved version of Benders's method.

BENDERS'S METHOD REVISITED

by

Egon Balas and Christian Bergthaller

1. Introduction

Consider the mixed integer program

$$(P)\qquad \begin{aligned} &\min cx + gy \\ &Ax + Dy = b \\ &x \geq 0,\ 0 \leq y \leq q, \\ &y \in \Omega \end{aligned}$$

and its linear programming relaxation (LP), obtained by removing the condition $y \in \Omega$. Here A is $m \times r$, D is $m \times n$, and $\Omega \subset R^n$ is an arbitrary finite set.

Benders [1] has shown that (P) is equivalent to (in the sense of having the same y-component for an optimal solution, as)

$$(P_1)\qquad \begin{aligned} &\min w_o \\ &w_o \geq (g - u^kD)y + u^kb \qquad & k \in S \\ &0 \geq \quad - v^kDy + v^kb \qquad & k \in T \\ &q \geq y \geq 0\ ,\ y \in \Omega \end{aligned}$$

where S and T are the index sets for the extreme points and extreme direction vectors, respectively, of

$$U = \{u \in R^m | uA \leq c\}\ .$$

For any $y \in R^n$, consider the pair of dual linear programs

(LP(y)) $\min\{cx \mid Ax = b - Dy,\ x \geq 0\}$

and

(LD(y)) $\max\{u(b - Dy) \mid uA \leq c\}$.

The standard procedure (also due to Benders) for solving (P) by using the above equivalence is to consider a relaxation $(\overline{P}_1)$ of (P_1), which consists of minimizing w_o subject to $q \geq y \geq 0$, $y \in \Omega$, and <u>some</u> of the constraints indexed by S and T. At the start, the only constraints of $(\overline{P}_1)$ may be $q \geq y \geq 0$, $y \in \Omega$. A sequence of the following two steps is then applied.

1. Solve the current $(\overline{P}_1)$. Let (w_o^k, y^k) be the optimal solution obtained. Go to 2.

2. Solve $(LD(y^k))$. Let u^k be an optimal extreme point of U, if one exists; or else, let u^k be an extreme point, and v^k an extreme direction vector, such that $u^k + \lambda v^k \in U$, $\forall\ \lambda > 0$, and $v^k(b - Dy^k) > 0$. In both cases, u^k defines for $(\overline{P}_1)$ a constraint of the type indexed by S, while in the second case, v^k also defines a constraint of the type indexed by T. Add these constraints to $(\overline{P}_1)$ and go to 1.

At every iteration, the minimum w_o^k of the current $(\overline{P}_1)$ provides a lower bound on the value of an optimal solution to (P_1), while $u^k(b - Dy^k) + gy^k$ obtained from solving $(LD(y^k))$, provides an upper bound. The lower bound w_o^k is monotone increasing. The procedure stops when the upper and lower bounds become equal.

The main difficulty with the above procedure is that in order to generate the subset of inequalities of (P_1) required to identify an optimal y, one has to repeatedly solve problems of the form $(\bar{P}_1)$, a computationally difficult task. It is therefore of interest to find other ways of generating constraints for $(\bar{P}_1)$. In one such attempt, D. McDaniel and M. Devine [2] have temporarily removed the constraint $y \in \Omega$ from $(\bar{P}_1)$ in the above two-step procedure, i.e., have temporarily replaced $(\bar{P}_1)$ by its linear programming relaxation. This change amounts to applying Benders's procedure to (LP) instead of (P). In the process of solving (LP) by Benders's procedure, a subset of the inequalities of (P_1) is generated. Furthermore, it was found that using these inequalities to define the initial problem $(\bar{P}_1)$ in Benders's procedure as applied to (P), has resulted in finding an optimal solution to (P) in a few iterations, often just one [2].

This suggests that the set of those inequalities of (P_1) that are tight for $(\bar{w}_o, \bar{y})$, where $(\bar{x}, \bar{y})$ is an optimal solution to (LP) and $\bar{w}_o = c\bar{x} + g\bar{y}$, or some appropriate subset of this set, is a highly desirable starting point for Benders's procedure, and may in fact yield an optimal solution to (P) in one or two iterations. The index set for these inequalities will be denoted by $S(\bar{y})$, i.e., we define

$$S(\bar{y}) = \{k \in S \mid \bar{w}_o = (g - u^k D)\bar{y} + u^k b\}.$$

In this paper we describe a new version of Benders's procedure, which is a considerable improvement over the original one. First, we characterize the subsystem of those inequalities of (P_1) indexed by $S(\bar{y})$, in terms of the simplex tableau associated with the optimal solution $(\bar{x}, \bar{y})$ to (LP). The cardinality of $S(\bar{y})$ is bounded by 2^p, where p is the number of basic components of $\bar{y}$. This is a best possible bound, which is attained quite frequently. Thus the optimal solution $(\bar{w}_o, \bar{y})$ to (LP_1), the linear programming relaxation of (P_1), is usually highly degenerate (2^p is usually considerably larger than $n + 1$, the number of inequalities that have to be satisfied with equality by any basic solution). Nevertheless, we show that none of the inequalities indexed by $S(\bar{y})$ is redundant, in the sense of being implied by the remaining inequalities of (P_1). Thus each of these inequalities may be needed to define an optimal solution to (P_1), though only $p + 1$ inequalities are needed to define, together with the inequalities $y_i \geq 0$ or $y \leq q_i$ for the nonbasic components of $\bar{y}$, the optimal solution $(\bar{w}_o, \bar{y})$ to (LP_1). We give a procedure which generates as many of the inequalities indexed by $S(\bar{y})$ as desired, at the cost of one pivot for each inequality, except for the first one, which requires p pivots. The improved Benders algorithm then consists of first using the above procedure to generate an initial constraint set for (P_1), namely an appropriate subset of the set indexed by $S(\bar{y})$, and then continue as usual. Computational experiments to determine the optimal number of initial constraints to be generated are under way.

2. The Binding Constraints of (LP_1)

An optimal basic solution $(\bar{x}, \bar{y})$ to (LP), with the basic and nonbasic components of x and y indexed by I_x, J_x and I_y, J_y respectively, can be represented in simplex tableau format as

$$z = a_{oo} + \sum_{j \in J_x} a_{oj}(-x_j) + \sum_{j \in J_y} a_{oj}(-y_j)$$

$$(1) \qquad x_i = a_{io} + \sum_{j \in J_x} a_{ij}(-x_j) + \sum_{j \in J_y} a_{ij}(-y_j) , \qquad i \in I_x ,$$

$$y_i = a_{io} + \sum_{j \in J_x} a_{ij}(-x_j) + \sum_{j \in J_y} a_{ij}(-y_i) , \qquad i \in I_y ,$$

where $a_{oj} \leq 0$, $\forall\, j \in J_x$, and where $(\bar{x}, \bar{y})$ is defined by

$$\bar{x}_i = \begin{cases} a_{io} & i \in I_x \\ 0 & i \in J_x \end{cases}$$

and

$$\bar{y}_i = \begin{cases} a_{io} & i \in I_y \\ 0 & i \in J_y ,\ a_{oi} \leq 0 \\ q_i & i \in J_y ,\ a_{oi} > 0 \end{cases}$$

Here $a_{oo} = c\bar{x} + g\bar{y}$, the value of the optimal solution $(\bar{x}, \bar{y})$, to (LP). Note that $\bar{x}$ is an optimal solution to $(LP(\bar{y}))$, while $(\bar{w}_o, \bar{y})$, where $\bar{w}_o = a_{oo}$, is an optimal solution to (LP_1), the linear programming relaxation of (P_1) (i.e., the problem obtained from (P_1) by removing the condition $y \in \Omega$).

We will assume that A is of full row rank. Whenever this is not the case, unit vectors corresponding to artificial variables with appropriate costs can be introduced in order to make the assumption hold.

Now consider $\bar{x}$. As a (basic) optimal solution to $(LP(\bar{y}))$, $\bar{x}$ is clearly degenerate, since substituting $\bar{y}$ for y in the righthand side vector of $(LP(y))$ sets all the basic variables corresponding to the rows indexed by I_y in tableau (1) equal to zero. Thus there is more than one basis that can be associated with $\bar{x}$. Actually, every basis obtained from (1) by a sequence of $|I_y|$ pivots, each of which replaces some y_i, $i \epsilon I_y$, with some x_j, $j \epsilon J_x$, produces an optimal basic solution to $(LP(\bar{y}))$. Not every such basis, however, corresponds to a feasible solution to $(LD(\bar{y}))$, i.e., to a point $u \epsilon U$. Those bases associated with $\bar{x}$ that correspond to feasible solutions u^k to $(LD(\bar{y}))$, i.e., to (extreme) points u^k of U, define the inequalities of (P_1) indexed by $S(\bar{y})$. These are the inequalities that we wish to generate.

Let $R = \{1,\ldots,r\}$ and $N = \{1,\ldots,n\}$ be the index sets associated with x and y respectively, and let $N_1 = \{j \epsilon N | \bar{y}_j = 1\}$.

Consider a sequence of simplex tableaus and pivots defined by the following rule. Let the k^{th} tableau be

$$z = a^k_{oo} + \sum_{j \epsilon J^k_x} a^k_{oj}(-x_j) + \sum_{j \epsilon J^k_y} a^k_{oj}(-y_j)$$

$$x_i = a^k_{io} + \sum_{j \epsilon J^k_x} a^k_{ij}(-x_j) + \sum_{j \epsilon J^k_y} a^k_{ij}(-y_j)\ , \quad i \epsilon I^k_x \tag{2}$$

$$y_i = a^k_{io} + \sum_{j \epsilon J^k_x} a^k_{ij}(-x_j) + \sum_{j \epsilon J^k_y} a^k_{ij}(-y_j)\ , \quad i \epsilon I^k_y$$

where the starting tableau, corresponding to $k = 0$, is obtained from (1) by setting $a^o_{ij} = a_{ij}$, $\forall$ i, j, $j \neq 0$; $I^o_z = I_z$ and $J^o_z = J_z$ for z = x, y; and

$$a^o_{io} = a_{io} + \sum_{j \in N_1} a_{oj} q_j \quad , \quad i \in I^o_x \cup I^o_y \cup \{0\}.$$

The pivoting rule for the k^{th} tableau is to choose $i_* \in I^k_y$ and pivot either on $a_{i_* j_1}$, where j_1 is defined by

$$\text{(3)} \qquad \frac{|a^k_{oj_1}|}{|a^k_{i_* j_1}|} = \min_{j \in J^k_x | a^k_{i_* j} > 0} \frac{|a^k_{oj}|}{|a^k_{i_* j}|} ,$$

or on $a_{i_* j_2}$, where j_2 is defined by

$$\text{(4)} \qquad \frac{|a^k_{oj_2}|}{|a^k_{i_* j_2}|} = \min_{j \in J^k_x | a^k_{i_* j} < 0} \frac{|a^k_{oj}|}{|a^k_{i_* j}|} .$$

Note that upper bounds don't play any role in this rule: y_{i_*} is always pivoted out of the basis at its lower bound of 0, and when the pivoting occurs on $a_{i_* j_2} < 0$, then x_{j_2} enters the basis with a negative value, i.e., is decreased rather than increased from its lower bound of 0.

<u>Theorem 1</u>. For any sequence of k pivots following the rule defined by (3), (4), where $0 \leq k \leq |I_y^o|$,

$$(5) \qquad w_o + \sum_{j \in J_y^k} a_{oj}^k y_j \geq a_{oo}^k$$

is a valid inequality for (P_1), which is satisfied with equality by $(\bar{w}_o, \bar{y})$.

<u>Proof</u>. Let a_j and d_j denote the j^{th} column of A and D respectively, and let $\gamma = (c, g)$. If B_k denotes the basis associated with the k^{th} tableau (2) and γ_{B_k} stands for the vector of basic components of γ, then in the starting tableau (k = 0) we have

$$a_{oj}^o = \gamma_{B_o} B_o^{-1} a_j - c_j \leq 0 \ , \qquad j \epsilon J_x^o$$

and

$$\gamma_{B_o} B_o^{-1} a_j - c_j = 0, \qquad j \epsilon I_x^o \ .$$

Therefore $u^o = \gamma_{B_o} B_o^{-1}$ is a feasible solution to $(LD(\bar{y}))$.

Further, the pivoting rules defined by (3) and (4) preserve the signs of the coefficients $a_{oj}^k = \gamma_{B_k} B_k^{-1} a_j - c_j$, $j \epsilon J_x^k$, and of course $\gamma_{B_k} B_k^{-1} a_j - c_j = 0$ holds for all $j \epsilon I_x^k$ after each pivot.

Hence each vector $u^k = \gamma_{B_k} B_k^{-1}$ is a feasible solution to $(LD(\bar{y}))$.

Further, we have

$$a_{oj}^k = u^k d_j - g_j \ , \qquad \forall \ j \epsilon J_y^k \ ,$$

and

$$u^k d_j - g_j = 0 \ , \qquad \forall \ j \epsilon I_y^k \ .$$

On the other hand, for each k we have

$$a_{oo}^k = \gamma_{B_k} B_k^{-1} b = u^k b \ .$$

Thus each inequality (5) can be restated as

$$w_o \geq (g - u^k D)y + u^k b ,$$

where $u^k \epsilon U$; hence each such inequality is valid for (P_1). To show that it is satisfied with equality by $(\bar{w}_o, \bar{y})$, we note that while each $u^k = \gamma_B B_k^{-1}$ is a feasible solution to $(LD(\bar{y}))$, it also satisfies the complementary slackness condition

$$(c - u^k A)x = 0$$

for $x = \bar{x}$, which is a feasible (and optimal) solution to $(LP(\bar{y}))$. Hence each u^k is an optimal solution to $(LD(\bar{y}))$, and therefore

$$u^k(b - D\bar{y}) = c\bar{x}$$

and adding to both sides $g\bar{y}$ produces

$$\bar{w}_o = c\bar{x} + g\bar{y}$$

$$= (g - u^k D)\bar{y} + u^k b \ . \qquad \underline{\text{Q.E.D.}}$$

While each of the inequalities (5) is valid for (P_1) and satisfied with equality by $(\bar{w}_o, \bar{y})$, these inequalities do not usually belong to the set indexed by $S(\bar{y})$ (or, for that matter, by S) except for the case when $k = |I_y^o|$, i. e., when $I_y^k = \emptyset$. The reason for this is that the vectors $u^k = \gamma_{B_k} B_k^{-1}$ obtained by fewer than $|I_y^o|$ pivots are usually

<u>nonbasic</u> feasible solutions to $(LD(\bar{y}))$. To see this, note that each u^k may have as many as $m + |J_x^k|$ positive components, whereas a basic u^k is restricted to at most r such components; but for $k < |I_y^o|$, $|J_x^k| > r - m$. On the other hand, as the next theorem shows, all the inequalities indexed by $S(\bar{y})$ can be obtained from (1) by some sequence of $|I_y^o|$ pivots of the above type.

<u>Theorem 2</u>. The constraints of (P_1) indexed by $S(\bar{y})$ are precisely those inequalities (5) such that $J_y^k = N$, and each one of them can be obtained from the system (1) by some sequence of $|I_y^o|$ pivots of the kind defined by (3), (4). Further, none of these inequalities is implied by the other inequalities of (P_1).

<u>Proof</u>. Since A is of full row rank, every basic component of y can be pivoted out of the basis in exchange for some component of x. When $J_y^k = N$, then $I_y^k = \emptyset$, and all the columns of B_k are columns of A, while $\gamma_{B_k} = c_{B_k}$. Then $u^k = c_{B_k} B_k^{-1}$ is a basic feasible solution to $(LD(\bar{y}))$, i.e., an extreme point of U. Since the inequality (5) associated with each such u^k was shown to hold with equality for $(\bar{w}_o, \bar{y})$, each such inequality belongs to the set indexed by $S(\bar{y})$. Conversely, every inequality indexed by $S(\bar{y})$ is defined by an optimal basic solution u^k to $(LD(\bar{y}))$, and all such solutions can be obtained from the system (1) by replacing the components of y in the basis with components of x, while preserving dual feasibility. But the rules (3), (4) are easily seen to exhaust the class of pivots by which this can be done.

$$(\rho - \sigma)\bar{y} \leq - \sigma q$$

or

$$\rho\bar{y} \leq -\sigma(q - \bar{y}) \leq 0$$

which contradicts $\rho \geq 0$. <u>Q.E.D.</u>

<u>Corollary 2.1</u>. $|S(\bar{y})| \leq 2^p$, where $p = |I_y^o|$. This is a best possible bound.

<u>Proof</u>. There are 2^p possible sequences of pivots of the types defined by (3) and (4). It is trivial to construct examples in which both types of pivots are possible at each step, since this is the usual case. The numerical example at the end of the paper illustrates the point. <u>Q.E.D.</u>

The fact that 2^p is a best bound on $|S(\bar{y})|$, which is attained more often than not, and yet none of the constraints indexed by $S(\bar{y})$ is redundant, reveals an interesting feature of problem (P_1). Since n-p inequalities of the form $y_i \geq 0$ or $y_i \leq q_i$, corresponding to the nonbasic components of $\bar{y}$, are satisfied with equality by $(\bar{w}_o, \bar{y})$, only p + 1 inequalities of the set indexed by $S(\bar{y})$ are needed to define the optimal solution $(\bar{w}_o, \bar{y})$ to (LP_1). However, since none of the inequalities of $S(\bar{y})$ are implied by the other constraints of (P_1), they may <u>all</u> be needed to define an optimal solution to (P_1). While this is certainly a possibility, empirical evidence indicates that in most instances the size of the set needed to identify an optimal solution to (P_1) is closer to p + 1 than to 2^p. Of course, some inequalities indexed by $S \setminus S(\bar{y})$ may also be needed.

The rule defined by (3), (4) can be used to obtain all basic optimal solutions to $(LD(\bar{y}))$, hence all the inequalities of (P_1) indexed by $S(\bar{y})$, by starting each time from tableau (1) and applying $|I_y^o|$ pivots.

However, after the first sequence of $|I_y^o|$ pivots, i.e., after obtaining the first basic optimal solution u^1 and an associated tableau (2) with $I_y^k = \emptyset$, there are better ways of finding additional optimal solutions, than by reverting to the starting tableau. A much cheaper procedure is as follows.

Let the last tableau (2), with $k = |I_y^o|$, be of the form

$$(2') \qquad \begin{aligned} z &= a_{oo}^k + \sum_{j \in J_x^k} a_{oj}^k(-x_j) + \sum_{j \in J_y^k} a_{oj}^k(-y_j) \\ x_i &= a_{io}^k + \sum_{j \in J_x^k} a_{ij}^k(-x_j) + \sum_{j \in J_y^k} a_{ij}^k(-y_j) \ , \ i \in I_x^k \ , \end{aligned}$$

with $J_y^k = N$. This is the same as (2), since for $k = |I_y^o|$, one has $I_y^k = \emptyset$. We recall that $a_{oj}^k \leq 0$, $\forall \ j \in J_x^k$. Consider now a pivot on $a_{i_* j_*}$, where $i_* \in I_x^k \setminus I_x^o$ is any row such that $a_{i_* j}^k < 0$ for at least one $j \in J_x^k$, and j_* is defined by

$$(9) \qquad \frac{|a_{oj_*}^k|}{|a_{i_* j_*}^k|} = \min_{j \in J_x^k | a_{i_* j}^k < 0} \frac{|a_{oj}^k|}{|a_{i_* j}^k|}$$

<u>Remark 1</u>. Any pivot in a tableau of the form (2′), based on the rule (9), produces a tableau of the same form, such that

$$(10) \qquad w_o + \sum_{j \in N} a_{oj}^k y_j \geq a_{oo}^k$$

is one of those inequalities of (P_1) indexed by $S(\bar{y})$.

<u>Proof</u>. The pivoting rule (9) replaces a basis B_k satisfying

$$(10) \quad \begin{aligned} &B_k^{-1}(b - D\bar{y}) \geq 0, \\ &c_{B_k} B_k^{-1} A - c \leq 0 \end{aligned}$$

with a basis B_{k+1} also satisfying (10) (with k replaced by k + 1). Indeed, the choice of $i_* \epsilon I_x^k \setminus I_x^o$ guarantees the first inequality of (10), while the choice of j_* by (9) guarantees the second one. Hence $u^{k+1} = c_{B_{k+1}} B_{k+1}^{-1}$ is a basic optimal solution to $(LD(\bar{y}))$. <u>Q.E.D</u>.

Note that, in order to obtain an inequality (10) from tableau (2′), one does not have to transform the whole tableau, but only the 0 row. In other words, a single tableau (2′) can serve for the computation of all the inequalities obtainable by exchanging any one of the basic variables x_i, $i \epsilon I_x^k \setminus I_x^o$, for the appropriate nonbasic variable x_j, $j \epsilon J_x^k$.

<u>3. An Improved Version of Benders's Algorithm</u>

Theorem 1 and Remark 1 provide a way of generating the inequalities of (P_1) indexed by $S(\bar{y})$ at the cost of one pivot for each new inequality except for the first one, which requires $|I_y^o|$ pivots.

As mentioned above, though the maximum number of inequalities obtainable from (1) is 2^p, the number actually needed to define a linear program whose set of optimal solutions is the same as that of (LP_1), is p + 1. This does not imply that <u>any</u> set of p + 1 inequalities of $S(\bar{y})$ defines such a linear program, only that such subsets of p + 1 inequalities exist. One can use certain devices to choose the inequalities that one generates so as to make it very likely that they belong to such a subset, but these devices have a computational cost, and even if they produce the desired result, it does not follow that the set of inequalities thus obtained is an adequate representation of (P_1) in the sense of having the same optimal solution (with $y \in \Omega$) as (P_1).

Therefore, in the algorithm described below, we chose the option of first generating t inequalities of $S(\bar{y})$ in a relatively easy way, where t is some integer satisfying $p + 1 \leq t \leq n + 1$, and then checking whether any additional inequalities are needed to define a linear program that adequately represents (LP_1). In case the test is negative, it also delivers one of the additional inequalities that are needed. This is repeated, if necessary, until an adequate representation of (LP_1) is obtained. At that point the problem $(\bar{P}_1)$ defined by the current set of inequalities (and the constraint $y \in \Omega$) is solved. The solution is then tested for optimality, and if the test fails, the procedure is continued as in the original Benders algorithm.

To simplify the exposition, we assume that $(LP(y))$ is feasible for all y generated during the procedure, i.e., that $T = \emptyset$ in (P_1). The extension to the general case is obvious.

Modified Benders Algorithm

Step 0. Solve (LP) by the simplex method for linear programs with bounded variables, and drive out of the basis every y_i which is at its lower or upper bound. Let the optimal solution obtained be $(\bar{x}, \bar{y})$, let $\bar{w}_o = c\bar{x} + g\bar{y}$, and let the associated simplex tableau be of the form (1). Make this into a tableau of the form (2), with $k = 0$, by replacing a_{io} with

$$a_{io}^{o} = a_{io} + \sum_{j \in N_1} a_{ij} q_j \quad , \quad i \in I_x^{o} \cup I_y^{o} \cup \{0\}.$$

Set t to the desired value $(p + 1 \leq t \leq n + 1)$, set $\beta = \infty$, and go to 1.

Step 1. If $I_y^k = \emptyset$, go to 2. Otherwise, choose $i_* \epsilon I_y^k$ and pivot either on $a^k_{i_* j_1}$ or on $a^k_{i_* j_2}$, where j_1 and j_2 are defined by

$$(3) \qquad \frac{|a^k_{oj_1}|}{|a^k_{i_* j_1}|} = \min_{j \epsilon J^k_x | a^k_{ij} > 0} \frac{|a^k_{oj}|}{|a^k_{i_* j}|} ,$$

and

$$(4) \qquad \frac{|a^k_{oj_2}|}{|a^k_{i_* j_2}|} = \min_{j \epsilon J^k_x | a^k_{i_* j} < 0} \frac{|a^k_{oj}|}{|a^k_{i_* j}|}$$

respectively. Then set $k \leftarrow k + 1$, and go to 1.

Step 2. Generate the inequality

$$w_o + \sum_{j \epsilon N} a^k_{oj} y_j \geq a^k_{oo} .$$

If the number of inequalities generated so far exceeds t, or $I^k_x \backslash I^o_x = \emptyset$, or $a^k_{ij} \geq 0$, $\forall\, j \epsilon J^k_x$, $\forall\, i \epsilon I^k_x \backslash I^o_x$, go to 3.

Otherwise, choose $i \epsilon I^k_x \backslash I^o_x$ such that $a^k_{ij} < 0$ for some $j \epsilon J^k_x$, pivot on $a^k_{ij_*}$ defined by

$$(9) \qquad \frac{|a^k_{oj_*}|}{|a^k_{ij_*}|} = \min_{j \epsilon J^k_x | a^k_{i_* j} < 0} \frac{|a^k_{oj}|}{|a^k_{i_* j}|} ,$$

and go to 2.

Step 3. Let W be the index set for the inequalities generated so far. Solve the problem

$$\min\{w_o \mid w_o + \sum_{j \in N} a^k_{oj} y_j \geq a^k_{oo},\ k \in W;\ 0 \leq y \leq q\},$$

and let (w_o^{k+1}, y^{k+1}) be the optimal solution found.

If $w_o^{k+1} < \bar{w}_o$, go to 4. Otherwise set $k \leftarrow k + 1$ and go to 3a.

Step 3a. This is Step 3 with the following changes:

(a) $y \in \Omega$ to be added to the constraints of the minimization problem

(b) $\bar{w}_o$ to be replaced by β

(c) If $w_o^{k+1} \geq \beta$, stop: y^k is optimal.

Step 4. Solve the linear program

$$(LD(y^{k+1})) \qquad \max\{u(b - Dy^{k+1}) \mid uA \leq c\}$$

(or its dual $(LP(y^{k+1}))$ and let u^{k+1} be the optimal solution found. Let $\beta \leftarrow \min\{\beta,\ u^{k+1}(b - Dy^{k+1}) + gy^{k+1}\}$.

If $\beta \leq w_o^{k+1}$ and Step 4 was entered from 3a, stop: (w_o^{k+1}, y^{k+1}) is optimal for (P_1).

If $\beta \leq w_o^{k+1}$ and Step 4 was entered from 3, then (w_o^{k+1}, y^{k+1}) is optimal for (LP_1). Set $\beta = \infty$, $k \leftarrow k + 1$ and go to 3a.

Otherwise, set $k \leftarrow k + 1$, generate the inequality

$$w_o + \sum_{j \in N} a^k_{oj} y_j \geq a^k_{oo},$$

where $a^k_{oo} = u^k b$ and a^k_{oj} is the j^{th} component of $(u^k D - g)$, and go to 3 (if Step 4 was entered from 3) or to 3a (if Step 4 was entered from 3a).

Discussion. At every iteration of step 1 a choice has to be made between the two types of pivots defined by (3) and (4). Since the strength of the cut obtained by a sequence of steps 1 grows with the size of the coefficients a^k_{oo} and $-a^k_{oj}$, $j \epsilon N$, it seems reasonable to choose the pivot which produces the greatest increase (smallest decrease) in some weighted sum of these coefficients, i.e., in

$$\alpha^k = a^k_{oo} + \sum_{j \epsilon J^k_y} \lambda_j(-a^k_{oj}).$$

As to the weights λ_j, one would like them to be proportional to probability of y_j being equal to q_j in an optimal solution. However, such information is usually not available. On the other hand, one may have some information on the number γ of variables y_j that have to be at their upper bounds in an optimal solution. If so, one can use $\lambda_j = \lambda = \gamma/|N|$, $\forall$ j. In the absence of any information of this type, a reasonable rule seems to be to use $\lambda_j = \lambda = 1/2$, $\forall$ j.

To implement the above choice rule, one can introduce a column

$$a^k_J = \sum_{j \epsilon J^k_y} a^k_j ,$$

with components a^k_{iJ}. Then the change in α^k as a result of the pivot at step k (i.e., the amount $\Delta = \alpha^{k+1} - \alpha^k$) is

$$\Delta(j_*) = (a^k_{i_*o} - \lambda a^k_{i_*J} - \lambda) \frac{(-a^k_{oj_*})}{a^k_{i_*j_*}}$$

for $j_* = j_1$ and $j_* = j_2$, and Δ can always be made nonnegative. To accomplish this, one chooses $j_* = j_1$ (i.e., $a^k_{i_*j_*} > 0$) if $a^k_{i_*o} - \lambda(a^k_{i_*J} + 1) \geq 0$, and $j_* = j_2$ (i.e., $a^k_{i_*j_*} < 0$) if $a^k_{i_*o} - \lambda(a^k_{i_*J} + 1) < 0$.

Similarly, in Step 2 one has to choose a row $i \in I_x^k \setminus I_x^o$. Again, a reasonable rule seems to be to maximize $\Delta = \alpha^{k+1} - \alpha^k$, which in this case can be expressed as a function of the index i (using the same column a_j^k as above);

$$\Delta(i) = (a_{io}^k - \lambda a_{ij}^k) \frac{(-a_{oj_i}^k)}{a_{ij_i}^k} .$$

Here j_i is the pivot column j_* prescribed by (9) for row i. One then pivots in the row i for which $\Delta(i)$ attains its maximum over $I_x^k \setminus I_x^o$, subject to some condition defined so as to prevent the repetition of bases (for instance, the condition may require all variables leaving the basis to remain nonbasic for a certain number of iterations).

4. Numerical Example

Consider the problem

$$\text{(P)} \qquad \begin{aligned} \min\ & x_1 + x_2 + 3x_3 + 5x_4 \qquad\qquad + 2y_7 - 4y_8 \\ & x_1 - x_2 + x_3 - 2x_4 + y_5 \qquad + 2y_7 - y_8 = 1 \\ & x_1 - 2x_2 - 8x_3 + 6x_4 \qquad + y_6 - y_7 + y_8 = 17 \\ & x_j \geq 0,\ j = 1,\dots,4;\ y_5, y_6 = 0 \text{ or } 15;\ y_7, y_8 = 0 \text{ or } 5 \end{aligned}$$

Step 0.

The optimal solution to (LP), the linear program obtained from (P) by replacing "y_5, y_6 = 0 or 15, y_7, y_8 = 0 or 5" by "$0 \leq y_5, y_6 \leq 15$, $0 \leq y_7, y_8 \leq 5$," is $\bar{x}_j = 0$, $j = 1,\ldots,4$; $\bar{y}_5 = 6$, $\bar{y}_6 = 12$, $\bar{y}_7 = 0$, $\bar{y}_8 = 5$, with the associated simplex tableau 1, where $N_1 = \{8\}$ (the index set for the nonbasic components of y at their upper bound).

	1	$-x_1$	$-x_2$	$-x_3$	$-x_4$	$-y_7$	$-y_8$	
z	-20	-1	-1	-3	-5	-2	4	$\bar{w}_o = c\bar{x} + g\bar{y} = -20$
y_5	6	$\boxed{1}$	-1	1	-2	2	-1	
y_6	12	1	-2	-8	6	-1	1	$y_8 = 5$

Tableau 1.

We replace a_{io} by $a^o_{io} = a_{io} + 5a_{i5}$ for $i = 0, 5, 6$, and form the column $a^o_J = a^o_7 + a^o_8$. This produces tableau 2.

	1	$-x_1$	$-x_2$	$-x_3$	$-x_4$	$-y_7$	$-y_8$	$-y_J$
z	0	-1	-1	-3	-5	-2	4	2
y_5	1	$\boxed{1}$	-1	1	-2	2	-1	1
y_6	17	1	-2	-8	6	-1	1	0

Tableau 2.

have $p = 2$, $n = 4$; and we set $t = 4$ $(p + 1 \leq 4 \leq n + 1)$.

Step 1. We choose $i_* = 5$. For choosing the pivot type, we use the rule discussed at the end of section 3, with $\lambda_j = \lambda = \frac{1}{2}$, $\forall j$.

Since $a^o_{50} - \frac{1}{2}(a^o_{5J} + 1) = 0$, we choose $j_* = j_1$, defined by (3). Pivoting on $a^o_{51} = 1$ and updating a_J (by adding the new non-basic column corresponding to y_5) yields tableau 3.

	1	$-y_5$	$-x_2$	$-x_3$	$-x_4$	$-y_7$	$-y_8$	$-y_J$
z	1	1	-2	-2	-7	0	3	4
x_1	1	1	-1	1	-2	2	-1	2
y_6	16	-1	-1	-9	$\boxed{8}$	-3	2	-2

Tableau 3.

Here $I^1_x = \{1\}$, $J^1_x = \{2,3,4\}$, $I^1_y = \{6\}$, $J^1_y = \{5,7,8\}$. Since $I^1_y \neq \emptyset$, we go to 1.

Step 1. We choose $i_* = 6$ and, since $a^1_{60} - \frac{1}{2}(a^1_{6J} + 1) = 31/2 > 0$, we again choose $j_* = j_1$. Pivoting on a^1_{68} and updating a_J then yields tableau 4.

	1	$-y_5$	$-x_2$	$-x_3$	$-y_6$	$-y_7$	$-y_8$	$-y_J$
z	15	$\frac{1}{8}$	$-\frac{23}{8}$	$-\frac{79}{8}$	$\frac{7}{8}$	$-\frac{21}{8}$	$\frac{38}{8}$	$\frac{25}{8}$
x_1	5	$\frac{6}{8}$	$\boxed{-\frac{10}{8}}$	$-\frac{10}{8}$	$\frac{2}{8}$	$\frac{10}{8}$	$-\frac{4}{8}$	$\frac{14}{8}$
x_4	2	$-\frac{1}{8}$	$-\frac{1}{8}$	$-\frac{9}{8}$	$\frac{1}{8}$	$-\frac{3}{8}$	$\frac{2}{8}$	$-\frac{1}{8}$

Tableau 4.

Since $I^2_y = \emptyset$, we go to 2.

Step 2. From tableau 4 we generate the inequality

$$w_o + \frac{1}{8} y_5 + \frac{7}{8} y_6 - \frac{21}{8} y_7 + \frac{19}{4} y_8 \geq 15.$$

We apply 3 more times Step 2 in order to obtain $t = 4$ cuts. The sequence of pivots and the cuts obtained by them are shown in tableaus 4-7.

	1	$-y_5$	$-x_1$	$-x_3$	$-y_6$	$-y_7$	$-y_8$
z	$\frac{7}{2}$	$-\frac{16}{10}$	$-\frac{23}{10}$	-7	$\frac{3}{10}$	$-\frac{11}{2}$	$\frac{59}{10}$
x_2	-4	$-\frac{6}{10}$	$-\frac{8}{10}$	1	$-\frac{2}{10}$	-1	$\frac{4}{10}$
x_4	$\frac{3}{2}$	$-\frac{2}{10}$	$-\frac{1}{10}$	$\boxed{-1}$	$\frac{1}{10}$	$-\frac{1}{2}$	$\frac{3}{10}$

Tableau 5.

$$w_o - \frac{8}{5} y_5 + \frac{3}{10} y_6 - \frac{11}{2} y_7 + \frac{59}{10} y_8 \geq \frac{7}{2}$$

	1	$-y_5$	$-x_1$	$-x_4$	$-y_6$	$-y_7$	$-y_8$
z	-7	$-\frac{2}{10}$	$-\frac{16}{10}$	-7	$-\frac{4}{10}$	-2	$\frac{38}{10}$
x_2	$-\frac{5}{2}$	$-\frac{8}{10}$	$\boxed{-\frac{9}{10}}$	1	$-\frac{1}{10}$	$-\frac{3}{2}$	$\frac{7}{10}$
x_3	$-\frac{3}{2}$	$\frac{2}{10}$	$\frac{1}{10}$	-1	$-\frac{1}{10}$	$\frac{1}{2}$	$-\frac{3}{10}$

Tableau 6.

$$w_o - \frac{1}{5} y_5 - \frac{2}{5} y_6 - 2y_7 + \frac{19}{5} y_8 \geq -7$$

	1	$-y_5$	$-x_2$	$-x_4$	$-y_6$	$-y_7$	$-y_8$
z	$-\frac{23}{9}$	$\frac{11}{9}$	$-\frac{16}{9}$	$-\frac{79}{9}$	$-\frac{2}{9}$	$\frac{6}{9}$	$\frac{23}{9}$
x_1	$\frac{25}{9}$	$\frac{8}{9}$	$-\frac{10}{9}$	$-\frac{10}{9}$	$\frac{1}{9}$	$\frac{15}{9}$	$-\frac{7}{9}$
x_3	$-\frac{16}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$-\frac{8}{9}$	$-\frac{1}{9}$	$\frac{3}{9}$	$-\frac{2}{9}$

Tableau 7.

$$w_o + \frac{11}{9} y_5 - \frac{2}{9} y_6 + \frac{2}{3} y_7 + \frac{23}{9} y_8 \geq -\frac{23}{9}$$

Since we now have t = 4 inequalities, we go to 3.

Step 3. We solve the linear program

min w_o

$$w_o + \frac{1}{8}y_5 + \frac{7}{8}y_6 - \frac{21}{8}y_7 + \frac{19}{4}y_8 \geq 15$$

$$w_o - \frac{8}{5}y_5 + \frac{3}{10}y_6 - \frac{11}{2}y_7 + \frac{59}{10}y_8 \geq \frac{7}{2}$$

$$w_o - \frac{1}{5}y_5 - \frac{2}{5}y_6 - 2y_7 + \frac{19}{5}y_8 \geq -7$$

$$w_o + \frac{11}{9}y_5 - \frac{2}{9}y_6 + \frac{2}{3}y_7 + \frac{23}{9}y_8 \geq -\frac{23}{9}$$

$$0 \leq y_5,\ y_6 \leq 15, \quad 0 \leq y_7,\ y_8 \leq 5$$

and find that its optimal solution is $(\hat{w}_o, \hat{y}) = (-20; 6, 12, 0, 5)$. Since $\hat{w}_o \geq \bar{w}_o = -20$, we go to 3a.

Step 3a. We solve the discrete programming problem obtained from the linear program of step 3 by requiring each component of y to be equal to one of its bounds, and find the optimal solution $(\tilde{w}_o, \tilde{y}) = (-12; 0, 15, 0, 5)$. We go to 4.

Step 4. We solve $(LP(\tilde{y}))$ rather than $(LD(\tilde{y}))$ by setting $y_1 = 0$, $y_2 = 15$, $y_3 = 0$, $y_4 = 5$ in Tableau 4, and reoptimizing the resulting linear program, as shown in Tableaus 5-6:

		$-x_1$	$-x_4$
z	-20	$-\frac{8}{5}$	-7
x_2	$-\frac{9}{2}$	$\boxed{-\frac{9}{10}}$	1
x_3	$\frac{3}{2}$	$\frac{1}{10}$	-1

Tableau 5.

		$-x_2$	$-x_4$
z	-12	$-\frac{16}{9}$	$-\frac{79}{9}$
x_1	5	$-\frac{10}{9}$	$-\frac{10}{9}$
x_3	1	$-\frac{1}{9}$	$-\frac{8}{9}$

Tableau 6.

Since the optimal solution $\tilde{x}$ gives $c\tilde{x} + g\tilde{y} = -12 \leq \tilde{w}_o$, $(\tilde{w}_o, \tilde{y})$ is an optimal solution to (P_1), and $(\tilde{x}, \tilde{y})$ is an optimal solution to (P), with $\tilde{x} = (5, 0, 1, 0)$, $\tilde{y} = (0, 15, 0, 5)$, and $c\tilde{x} + g\tilde{y} = -12$.

References

[1] J. F. Benders, "Partitioning Procedures for Solving Mixed-Variables Programming Problems," Numerische Mathematik, 4, 1962, 238-252.

[2] D. McDaniel and M. Devine, "A Modified Benders' Partitioning Algorithm for Mixed Integer Programming." University of Wisconsin at Milwaukee, 1975.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
Technical ~~Report~~ No. 401

2. GOVT ACCESSION NO. 3. RECIPIENT'S CATALOG NUMBER
(9) Management sciences research rept.,

4. TITLE (and Subtitle)
(6) Benders's Method Revisited,

5. TYPE OF REPORT & PERIOD COVERED
Technical Report
May 1977

6. PERFORMING ORG. REPORT NUMBER
(14) MSRR-~~401~~, WP-42-76-77

7. AUTHOR(s)
(10) Egon/Balas
Christian/Bergthaller

8. CONTRACT OR GRANT NUMBER(s)
(15) N00014-75-C-0621,
✓NSF-MPS73-08534 ~~A02 NSF~~

9. PERFORMING ORGANIZATION NAME AND ADDRESS
~~Graduate School of Industrial Admini~~stration
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
NR 047-048

11. CONTROLLING OFFICE NAME AND ADDRESS
Personnel and Training Research Programs
Office of Naval Research (Code 434)
Arlington, Virginia 22217

12. REPORT DATE
(11) May 1977

13. NUMBER OF PAGES
~~24~~ (12) 29p.

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

mixed integer programming, large scale systems, decomposition,
Benders's partitioning

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The master problem in Benders's partitioning method is an integer program with a very large number of constraints, each of which is usually generated by solving the integer program with the constraints generated earlier. Computational experience shows that the subset B of those constraints of the master problem that are satisfied with equality at the linear programming optimum often play a crucial role in determining the integer optimum, in the sense that only a few of the remaining inequalities are needed. We characterize this subset B of inequalities. Though the best upper bound — next page

(Cont'd)

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

403426 Jme

2 to the p power

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

(often attained) on the cardinality of B is 2^p, where p is the number of integer-constrained variables that are basic at the linear programming optimum, none of the inequalities in B is implied by the remaining inequalities of the master problem. We then give an efficient procedure for generating an appropriate subset of the inequalities in B, which leads to a considerably improved version of Benders's method.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

AD-A048 887 CARNEGIE-MELLON UNIV PITTSBURGH PA MANAGEMENT SCIENC--ETC F/G 12/1
BENDERS'S METHOD REVISITED.(U)
MAY 77 E BALAS, C BERGTHALLER N00014-75-C-0621
UNCLASSIFIED MSRR-401 NL

2 OF 2
AD
A048887

SUPPLEMENTARY
INFORMATION

END
DATE
FILMED
12-78
DDC

SUPPLEMENTARY

INFORMATION

Carnegie Univ., Pittsburgh, PA
Graduate school of Industrial Adm

ERRATA

for

Benders's Method Revisited

by

E. Balas and C. Bergthaller

The proof of the last statement in Theorem 2 (starting on top of page 11) has a flaw. The Theorem itself, however, is correct. Please substitute the attached proof.

To prove the last statement of the Theorem by contradiction, suppose that for some $h \in S(\bar{y})$ the inequality

$$w_o + \sum_{j \in N} a^h_{oj} y_j \geq a^h_{oo} ,$$

which can also be written as

$$(6) \qquad w_o + (u^h D - g)y \geq u^h b ,$$

is implied by the other inequalities of (P_1). Then

$$u^h b = \min\{w_o + (u^h D - g)y \mid w_o + (u^k D - g)y \geq u^k b,\ k \in S\setminus\{h\},\ 0 \leq y \leq q\},$$

and this minimum is attained for the same point $(\bar{w}_o, \bar{y})$ for which w_o attains its minimum in (P_1). Furthermore, the dual of the linear program in the brackets has an optimal solution (λ) satisfying

$$\sum_{k \in S-\{h\}} \lambda_k = 1,\ \lambda \geq 0;$$

$$\sum_{k \in S-\{h\}} \lambda_k (u^k d_j - g_j) \begin{cases} \leq u^h d_j - g_j, & \text{if } \bar{y}_j = 0 \\ = u^h d_j - g_j, & \text{if } 0 < \bar{y}_j < q_j \\ \geq u^h d_j - g_j, & \text{if } \bar{y}_j = q_j \end{cases} .$$

Denoting $N^+ = \{j \in N \mid 0 < \bar{y}_j < q_j\}$, from the above we have

$$(7) \qquad (\sum_{k \in S-\{h\}} \lambda_k u^k - u^h) d_j = 0,\ j \in N^+ .$$

Furthermore, any constraint with a positive multiplier λ_k must be tight at $(\bar{w}_o, \bar{y})$, i.e., $\lambda_k > 0$ implies $k \in S(\bar{y})$.

On the other hand, each u^k, $k \in S(\bar{y})$, is a basic optimal solution to $LD(\bar{y})$. Though the bases associated with the various u^k, $k \in S(\bar{y})$, differ among themselves, they all have among their basic columns those columns a_i of A which are basic in the optimal solution $(\bar{x},\bar{y})$ to (LP). If R^+ denotes the set of these columns, then $|R^+| = m - |N^+|$, where m is the number of rows of A (and of D). Thus, we have

$$u^k a_i = c_i \quad , \quad i \in R^+$$

for every $k \in S(\bar{y})$, and hence

$$(7') \qquad (\sum_{k \in S-\{h\}} \lambda_k u^i - u^h) a_i = 0, \quad i \in R^+.$$

Denoting

$$\delta = \sum_{k \in S-\{h\}} \lambda_k u_k - u^h,$$

equations (7) and (7′) can be restated as

$$(8) \qquad \begin{aligned} \delta d_j &= 0 \quad , \quad j \in N^+, \\ \delta a_i &= 0 \quad , \quad i \in R^+, \end{aligned}$$

where $|N^+| + |R^+| = m$, and where the m vectors d_j, a_i are linearly independent, since they are the columns of the basis associated with $(\bar{x},\bar{y})$.

Thus the unique solution of (8) is $\delta = 0$, and from (7) this implies that the vectors u^k, $k \in S(\bar{y})$ are linearly dependent, contrary to the assumption that they are extreme points of U. Q.E.D.